An Initial State Perturbation Experiment with the GISS Model[1,2]

JEROME SPAR AND JESUS J. NOTARIO

Department of Earth and Planetary Sciences, The City College, CUNY, New York, N. Y. 10031

WILLIAM J. QUIRK[3]

Goddard Institute for Space Studies, New York, N. Y. 10025

(Manuscript received 15 August 1977, in final form 7 October 1977)

ABSTRACT

Monthly mean global forecasts for January 1975 have been computed with the GISS model from four slightly different sets of initial conditions—a "control" state and three random perturbations thereof—to simulate the effects of initial state uncertainty on forecast quality. Differences among the forecasts are examined in terms of energetics, synoptic patterns and forecast statistics. The "noise level" of the model predictions is depicted on global maps of standard deviations of sea level pressures, 500 mb heights and 850 mb temperatures for the set of four forecasts. Initial small-scale random errors do not appear to result in any major degradation of the large-scale monthly mean forecast beyond that generated by the model itself, nor do they appear to represent the major source of large-scale forecast error.

1. Introduction

Because the state of the atmosphere can never be known exactly, numerical weather predictions will always contain inherent errors due to the uncertainty of initial conditions, along with the errors due to all other causes. As part of a continuing experiment in monthly mean weather prediction (Spar *et al.*, 1976; Spar, 1977) with the GISS model of the global atmosphere (Somerville *et al.*, 1974), we have attempted to evaluate the influence of random errors in the initial state on model-generated monthly mean prognostic fields.

Experiments with dynamical models of the atmosphere (e.g., National Academy of Sciences, 1966; Smagorinsky, 1969) have demonstrated that daily predictability will be effectively lost after two to three weeks as a result of initial state errors alone. The current operational forecast skill of dynamical models degrades even faster, being no better than climatology after one week (Miyakoda *et al.*, 1972). Time-averaging of model predictions, which has been done for periods of one to two weeks (Druyan *et al.*, 1975; Vanderman *et al.*, 1976) as well as one month (Spar *et al.*, 1976; Spar, 1977), appears to result in retention of some predictive skill. However, even monthly averages are sensitive to unavoidable errors in the initial state, as shown, for example, by the "noise level" experiments of Chervin and Schneider (1976a,b) with the NCAR model.

The noise level of a model may be evaluated by running it repeatedly with the same synoptic-scale initial conditions, but with the initial analysis contaminated for each forecast run by a different small-scale random perturbation field representing the inherent uncertainty of the initial state. The dispersion of the resulting forecasts represents the noise of the model. In the present study, the noise has been computed for the case of a monthly mean forecast for January 1975 generated by

TABLE 1. (A) Forecast and observed mean zonal available potential energy P_M and kinetic energy K_M (both in units of 10^5 J m^{-2}) over the Northern Hemisphere for January 1975. F, P-1, P-2 and P-3 denote the control and three perturbation forecasts, and O represents the observed atmosphere.

	F	P-1	P-2	P-3	O
P_M	60.4	61.4	61.1	60.0	55.5
K_M	8.9	9.2	8.8	9.0	7.8

(B) Forecast and observed maximum mean zonal wind speeds (m s^{-1}) (all at the 175 mb level of the model) and jet stream latitude in the Northern Hemisphere for January 1975.

	F	P-1	P-2	P-3	O
Maximum wind	37.1	39.7	39.0	40.0	33.1
Jet latitude (°N)	34.0	34.0	34.0	34.0	40.0

[1] This research was supported by the National Aeronautics and Space Administration, Goddard Space Flight Center, under Grant NGR 33-013-086, and was conducted largely at the Goddard Institute for Space Studies (GISS).

[2] Contribution No. 104, CUNY Institute of Marine and Atmospheric Sciences.

[3] Current affiliation: McKinsey and Co., Inc., 245 Park Ave., New York, N. Y. 10017

0027-0644/78/0089-0100$06.00

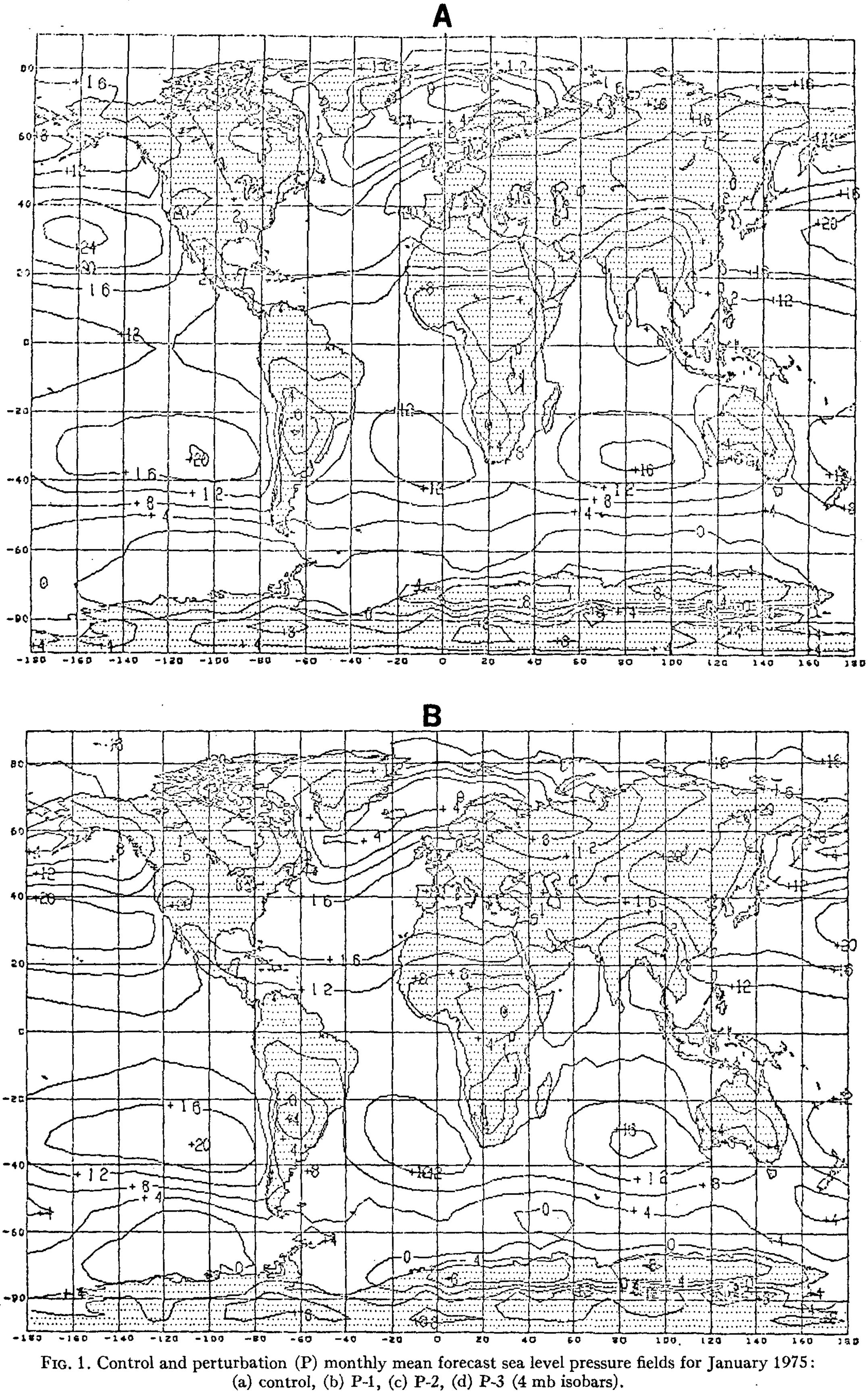

FIG. 1. Control and perturbation (P) monthly mean forecast sea level pressure fields for January 1975: (a) control, (b) P-1, (c) P-2, (d) P-3 (4 mb isobars).

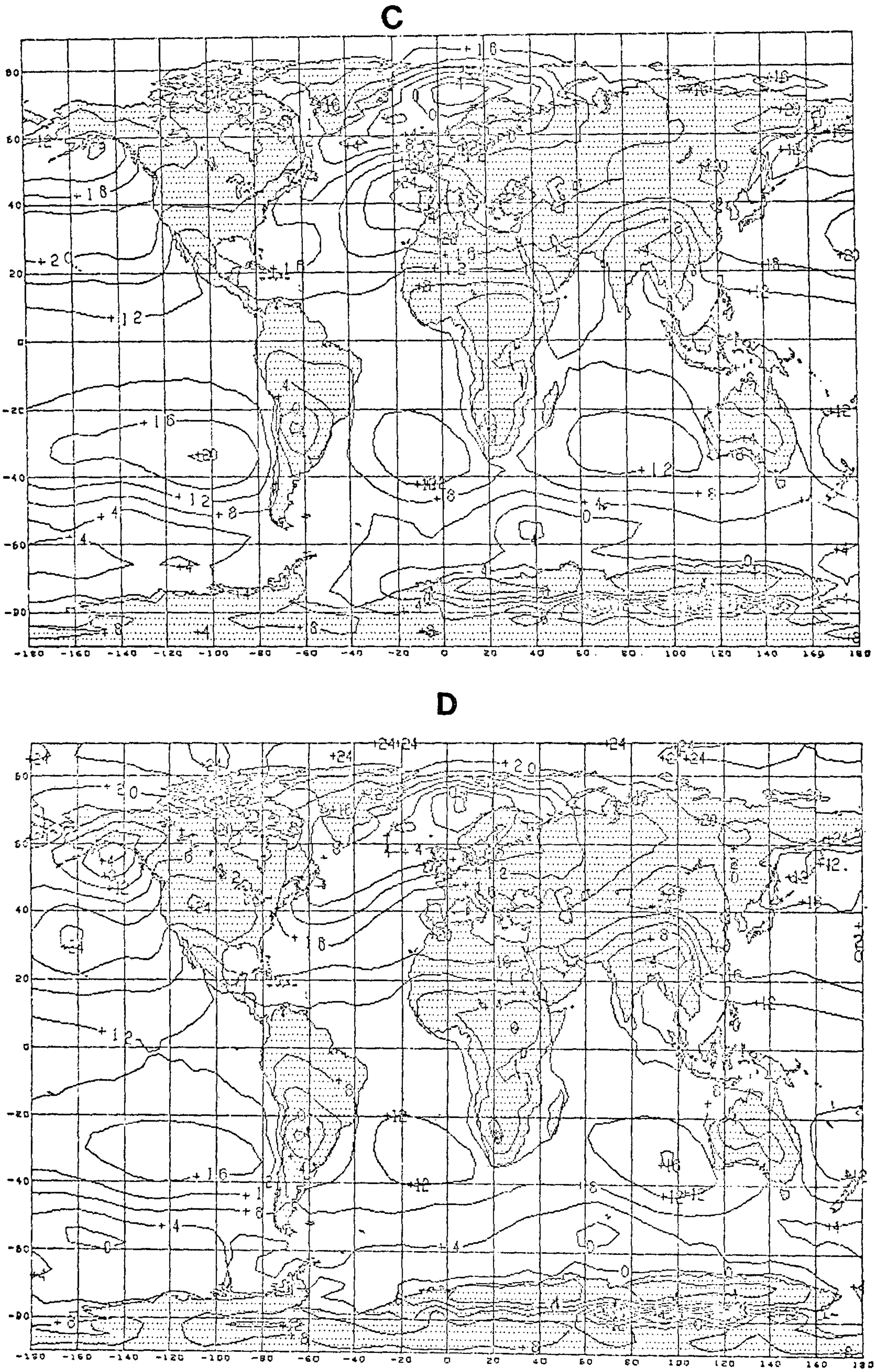

FIG. 1. *Continued*

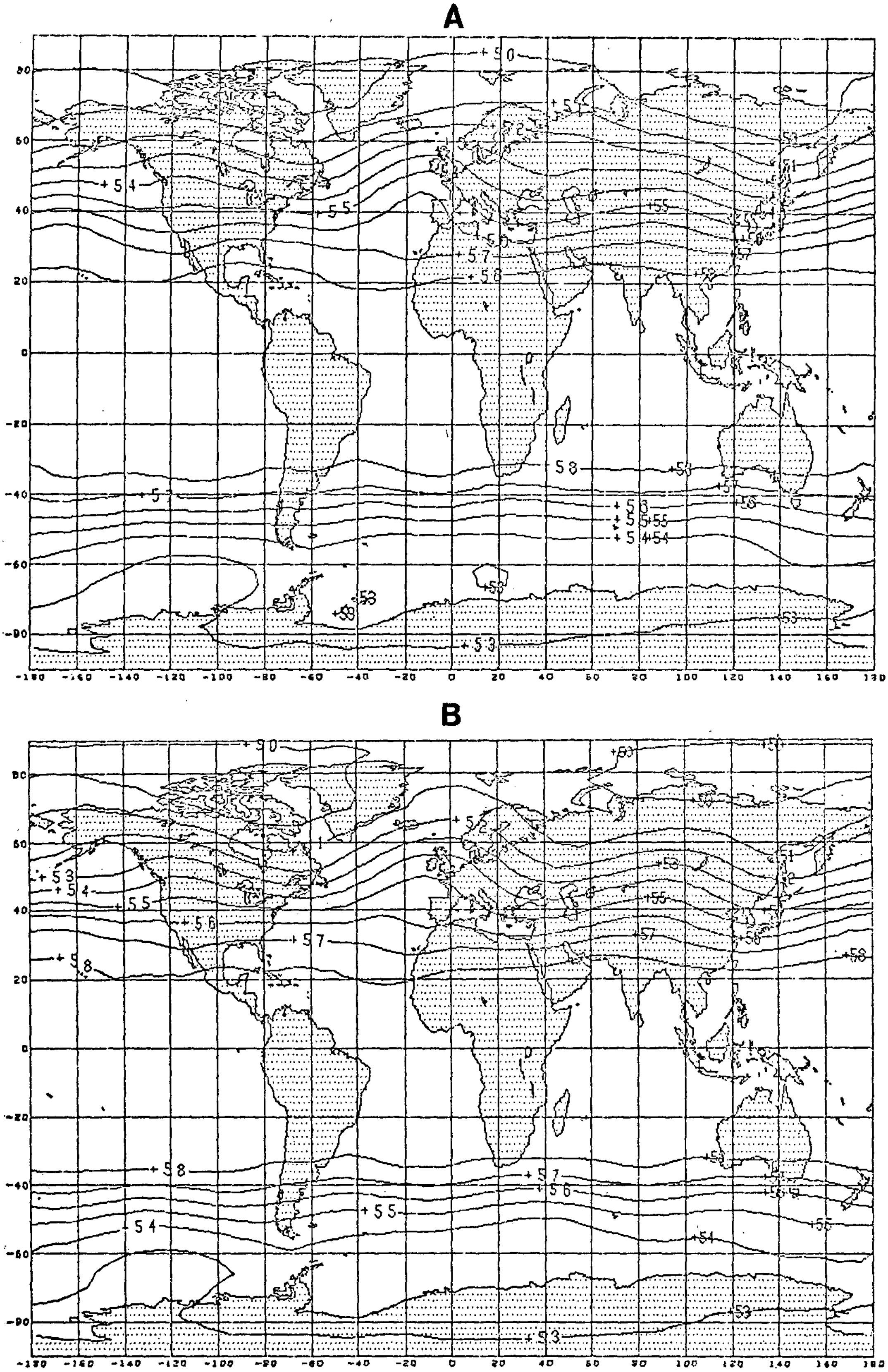

FIG. 2. Control and perturbation monthly mean forecast 500 mb height fields for January 1975: (a) Control, (b) P-1, (c) P-2, (d) P-3 (100 m contours).

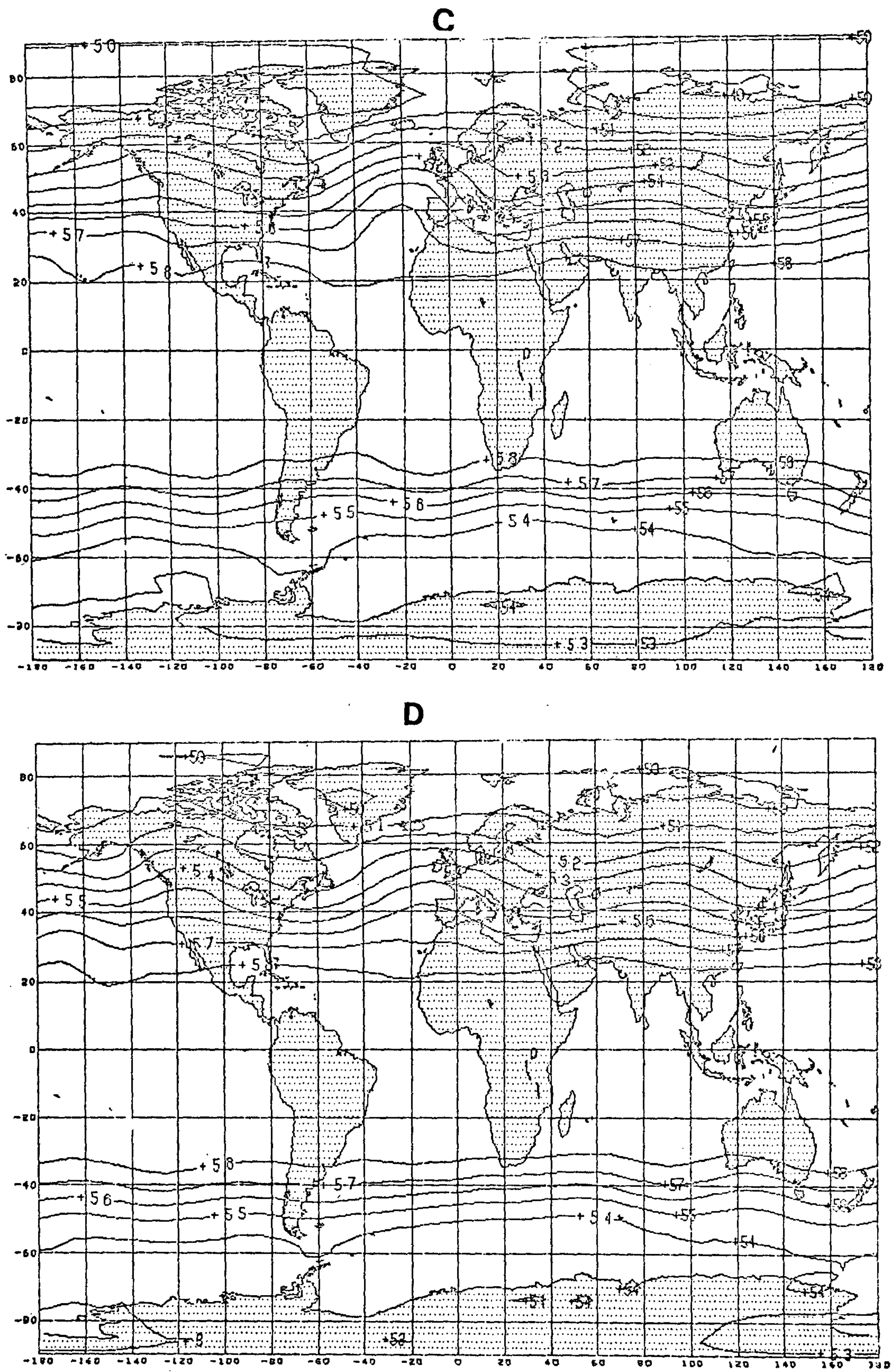

FIG. 2. *Continued*

the GISS model from initial conditions on the first day of that month. The principal objective of this experiment was to assess the effect of random initial state errors on such monthly mean forecasts.

2. Perturbation experiment

The basic initial state for the experiment was the 0000 GMT global analysis for 1 January 1975 provided by the National Meteorological Center (NMC). The error fields superimposed on this "control" state were produced by a random number generator which distributes the initial errors geographically. Each error set constitutes a Gaussian distribution with the same root-mean-square (rms) error for each of the variables perturbed. However, in the GISS perturbation experiments the rms errors were arbitrarily doubled over the the ocean compared with the land values in order to simulate poorer data coverage over water. Thus, the global rms errors inserted in the zonal and meridional wind components at all nine (sigma-coordinate) levels of the model are 4 m s^{-1} over land and 8 m s^{-1} over water. The rms errors in temperature inserted at all nine levels are 1 K over land and 2 K over water, while the rms errors in surface pressure are 3 mb over land and 6 mb over water.

The model computes various fields of interpolated and extrapolated quantities, both initially and during the forecast cycle. Thus, the initial computed rms errors of sea level pressure, 500 mb height and 850 mb temperature, for example, over the Northern Hemisphere in the experiment are approximately 2 mb, 18 m and 1 K, repectively, indicating considerable smoothing of the error field in the interpolation process.

Three different perturbations of the control initial state were generated, all with the same magnitudes for the global rms errors. The initial rms differences between any two perturbations are generally larger than the initial rms deviations from the control. They were, in fact, approximately 6 mb, 45 m and 2 K, respectively, for sea-level pressure, 500 mb height and 850 mb temperature over the Northern Hemisphere. However, it was found that 12 h into the forecast cycle the rms differences in sea level pressure and 500 mb height between any two perturbation runs decayed to very nearly the values of the corresponding initial rms errors. This fast adjustment was then followed by a relatively slow growth of the rms differences between forecasts, as compared with the rapid growth of forecast error measured against observation.

3. Results

The influence of random errors in initial conditions on the monthly mean forecasts may be examined, first of all, in terms of certain gross properties of the atmosphere. Table 1, for example, shows the mean zonal available potential energy (P_M) and mean zonal kinetic energy (K_M) over the Northern Hemisphere for both the four forecasts and the "observed" mean January 1975 atmosphere, the latter based on 12 h global NMC analyses. As in previous studies, the energies are integrated over the first eight layers of the GISS nine-layer model, i.e., up to about the 120 mb level. [For further details on the model energetics see Somerville *et al.* (1974), Tenenbaum (1976), Spar *et al.* (1976) and Spar (1977.)] Also shown in Table 1 are the maximum mean zonal wind speeds, both forecast and observed, all of which are found at the 175 mb level of the model, as well as the latitude of the maximum westerlies, i.e., the mean jet stream axis. The observed data have been reduced to the model grid and model resolution for comparability with the forecast profiles.

The predicted zonal energies in Table 1A lie within a relatively narrow range of values compared with the observed. Thus, all four potential energies forecast lie within 1% of the mean of the forecasts, compared with a mean error of about 9%. Similarly, the four kinetic energies forecast lie within 2% of the mean of the forecasts, compared with a mean error of 13%. The control forecast is not consistently better than the perturbation forecasts, and all four forecasts overpredict both the zonal available potential and zonal kinetic energy, reflecting a common tendency of the model to forecast an excessive meridional temperature gradient.

As shown in Table 1B, the four forecasts yield the same latitude (34°N) for the jet stream axis, and all commit the same error relative to the observed jet axis (40°N). On the other hand, the range of maximum wind speeds forecast, from 37 to 40 m s^{-1}, indicates a somewhat greater impact of the random errors on that parameter, and the control forecast (37 m s^{-1}) does lie closest to the observed value of 33 m s^{-1}. Nevertheless, the four forecasts are closer to each other than they are to the observed state.

As a further illustration of the role played by random initial errors in the monthly mean forecasts computed with the model, the four prognostic monthly mean global sea level pressure maps for January 1975 are shown in Fig. 1. There are several obvious differences among the forecast maps, especially in the Northern Hemisphere, where the structures of the eastern Atlantic high-pressure cell, the western Atlantic trough and the Aleutian low vary markedly. Nevertheless, when compared with the observed sea level pressure field, shown in Fig. 3, all four forecasts display similar major forecast errors, particularly in the eastern North Pacific Ocean and in the vicinity of Iceland.

The corresponding prognostic maps for the 500 mb level, shown in Fig. 2, may also be compared with each other, as well as with the observed mean 500 mb map in Fig. 3. Again, significant differences are found among

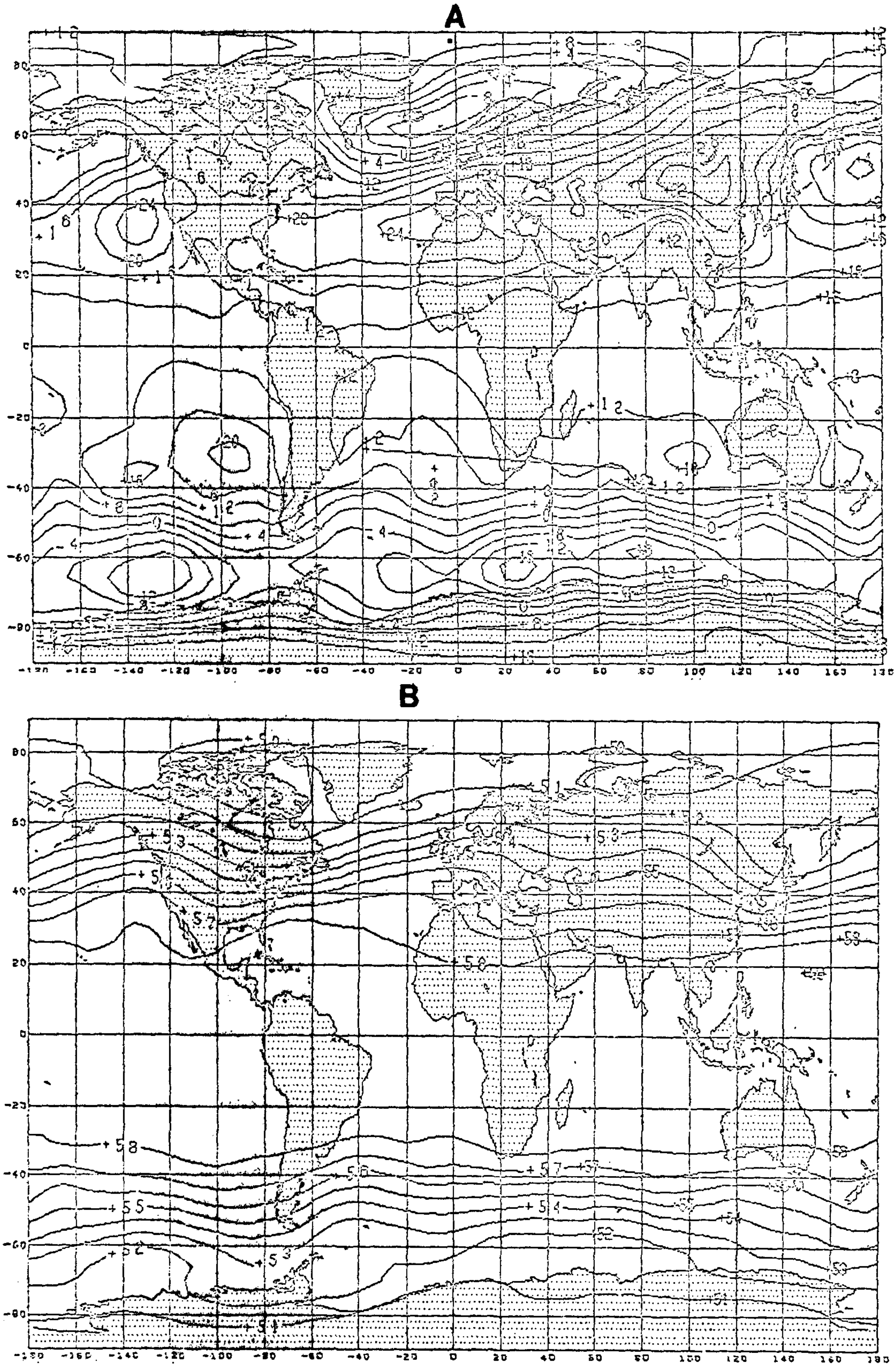

FIG. 3. Observed monthly mean sea level pressure (a) and 500 mb height (b) fields for January 1975, based on NMC analyses (4 mb isobars and 100 m contours).

TABLE 2. Root-mean-square errors and S_1 skill scores for the January 1975 control forecast (F), three perturbation forecasts (P-1, P-2, P-3), and a "forecast" of climatology (M).

(A) Sea level pressure	rms error (mb)					S_1 score				
Region	F	P-1	P-2	P-3	M	F	P-1	P-2	P-3	M
(1) Globe	5.1	5.3	5.5	5.5	5.9	67	67	69	69	72
(2) Northern Hemisphere	5.3	5.2	5.5	5.4	6.6	64	62	63	64	73
(3) Tropics	3.4	3.5	3.6	3.6	3.0	62	62	63	64	67
(4) E. Pacific U. S.	5.0	4.4	4.5	4.9	6.8					
(5) North America	4.2	3.6	4.6	3.9	10.0	86	78	89	91	97
(6) United States	3.5	2.9	4.0	3.4	7.1	85	79	91	87	103
(7) Europe	5.6	5.4	5.2	4.8	4.1	63	67	65	56	55

(B) 500 mb Height	rms error (m)					S_1 score				
Region	F	P-1	P-2	P-3	M	F	P-1	P-2	P-3	M
(1) Globe	60	61	61	64	68	42	40	41	43	49
(2) Northern Hemisphere	62	60	57	61	82	42	39	40	41	52
(3) Tropics	30	24	23	20	34	65	64	59	60	71
(4) E. Pacific U. S.	87	81	66	81	123					
(5) North America	60	45	56	70	130	38	32	33	41	50
(6) United States	58	47	51	55	117	41	35	35	42	40
(7) Europe	103	119	118	98	38	47	47	53	38	36

(C) 850 mb temperature	rms error (K)				
Region	F	P-1	P-2	P-3	M
(2) Northern Hemisphere	4.1	4.0	3.9	4.1	4.5
(4) E. Pacific U. S.	4.2	4.1	3.8	4.5	6.9
(6) United States	3.7	3.0	2.8	3.6	7.5

the Northern Hemisphere forecasts, notably in the ridge structures over the eastern Atlantic and over western North America. However, the major forecast error, which is the incorrect location of the North American trough, is repeated in all four forecasts.

The effects of random initial errors on the forecast statistics for January 1975 are shown in Table 2 for seven geographical regions. These include 1) the whole globe, 2) the Northern Hemisphere, 3) a tropical belt (22°N–22°S), 4) an east Pacific-United States belt (latitudes 30–54°N; longitudes 75°W–180°), 5) North America (latitudes 30–70°N; longitudes 75–130°W), 6) United States (latitudes 30–54°N; longitudes 75–130°W), and 7) Europe (latitudes 34–86°N; longitudes 10°W–40°E). Listed in Table 2 are rms errors and S_1 skill scores[4] (Teweles and Wobus, 1954) in sea-level pressure, 500 mb height and 850 mb temperature for the monthly mean forecasts for January 1975, generated from the control (F) and three perturbed (P-1, P-2, P-3) sets of initial conditions. Also tabulated are the corresponding scores for a "forecast" of climatology (M), in which the climatological monthly mean state for January is evaluated as a forecast for January 1975. [For further discussion of the model's error statistics, see Spar *et al.* (1976) and Spar (1977).]

With a few exceptions, the error scores of the control and perturbation forecasts shown in Table 2 lie within a relatively narrow range of values compared with those of the climatology forecast. Over the globe and also over the Northern Hemisphere, all four forecasts are superior to climatology, while over Europe all are worse. Again, the control forecast is not consistently better than the perturbation forecasts, and there is no evidence that random errors in the initial conditions resulted in any major degradation of the monthly mean forecasts beyond that generated by the model itself.

The differences among the monthly mean forecasts are further illustrated in Table 3, showing the rms differences and S_1 comparison scores[4] between forecasts P-1 and P-2 and between forecasts P-1 and P-3 in the same format as Table 2. From a comparison of the two tables it is evident that, with few exceptions, the rms differences between forecasts are considerably

[4] Due to a program oversight, S_1 scores were not computed for region (4), nor were they computed for the 850 mb temperatures.

smaller than the rms errors of the forecasts. In terms of S_1 scores, the differences between forecasts are again generally smaller than the forecast errors, the one outstanding exception being the sea level pressure field over Europe, where the differences exceed the forecast errors. Clearly, the differences among forecasts made from contaminated initial conditions are not insignificant. However, they are not large enough to suggest that random uncertainties in the initial state of the atmosphere represent the principal source of large-scale error in prognostic monthly mean maps.

[It is of interest to compare the rms differences and S_1 comparison scores in Table 3 with the results of one monthly mean forecast *replication* experiment that was carried out with the GISS model on a different data set in connection with another study, to be published separately. In that experiment, two monthly mean global forecasts were computed for February 1976 from identical initial and boundary conditions, with the same computer and identical programs, the two computations presumably differing only in the schedule of interruptions and restarts on the computer during the forecast history. Because of the cycling schedule of the model, and the fact that all calculations are not identical in each time step (Somerville *et al.*, 1974), it is apparently possible for small differences to develop between two long, interrupted forecast runs. Among the results of that experiment, it was found that, over the Northern Hemisphere, for example, the rms differences in sea level pressure, 500 mb height and 850 mb temperature were approximately 2 mb, 20 m and 1 K, respectively, between the two forecasts, while the S_1 comparison scores were approximately 40 and 20 for sea level pressure and 500 mb height, respectively. If these values are representative of the computational noise level of the monthly mean forecasts, there appears to be a minimal error resulting purely from the inability of the computer system to replicate perfectly a monthly mean forecast, and thus a large fraction of the difference between the perturbation forecasts indicated in Table 3 may be due to computational errors rather than initial state differences. [This, of course, does not include the additional computational uncertainty associated with the existence of alternative numerical approximations.]

Chervin and Schneider (1976b) have shown that there may be marked geographical variations in the sensitivity of model 30-day mean simulations to initial state random perturbations. To determine the uncertainties in the climatological behavior of the atmosphere as simulated by the NCAR model, they computed global maps of standard deviations of various 30-day mean fields, based on five perturbed model January simulations. In the NCAR experiment, the computations were started from an isothermal state of rest, and small perturbations were introduced on day 20 of a long simulation run. The 30-day averages were then computed over an interval (days 31–60, according to R. Chervin, personal communication) after transient predictability effects were no longer dominant. The five simulations are thus considered by the authors to be independent samples of a model January climatology.

TABLE 3. Root-mean-square differences and S_1 comparison scores for January 1975 between perturbation forecasts: P-1 vs P-2 and P-1 vs. P-3

(A) Sea level pressure

Region	rms difference (mb) P-1 vs P-2	rms difference (mb) P-1 vs P-3	S_1 comparison P-1 vs P-2	S_1 comparison P-1 vs P-3
(1) Globe	2.0	2.2	49	51
(2) Northern Hemisphere	2.3	2.5	53	54
(3) Tropics	0.8	0.8	40	40
(4) E. Pacific U. S.	2.4	2.9		
(5) North America	2.7	3.3	70	73
(6) United States	2.8	3.4	73	73
(7) Europe	3.5	3.3	72	64

(B) 500 mb height

Region	rms difference (m) P-1 vs P-2	rms difference (m) P-1 vs P-3	S_1 comparison P-1 vs P-2	S_1 comparison P-1 vs P-3
(1) Globe	25	25	28	29
(2) Northern Hemisphere	31	28	28	27
(3) Tropics	7	8	46	40
(4) E. Pacific U. S.	32	32		
(5) North America	30	47	25	28
(6) United States	26	35	25	25
(7) Europe	45	32	38	28

(C) 850 mb temperature

Region	rms difference (K) P-1 vs P-2	rms difference (K) P-1 vs P-3
(2) Northern Hemisphere	1.7	1.8
(4) E. Pacific U. S.	1.4	1.6
(6) United States	1.6	1.9

Standard deviations of monthly mean prognostic fields were also computed for the January 1975 GISS model forecasts, following a procedure similar to that of Chervin and Schneider (1976b). These are shown in Figs. 4–6 for sea level pressure, 500 mb height and 850 mb temperature, respectively. The NCAR and GISS experiments obviously differ in many important respects. The former used five simulations and four degrees of freedom for the estimates of the standard deviations, while only four forecasts and three degrees of freedom were used in the latter. The transient was discarded in the NCAR calculation and retained in the GISS experiment. Furthermore, the objectives and methodologies of the two studies are clearly quite

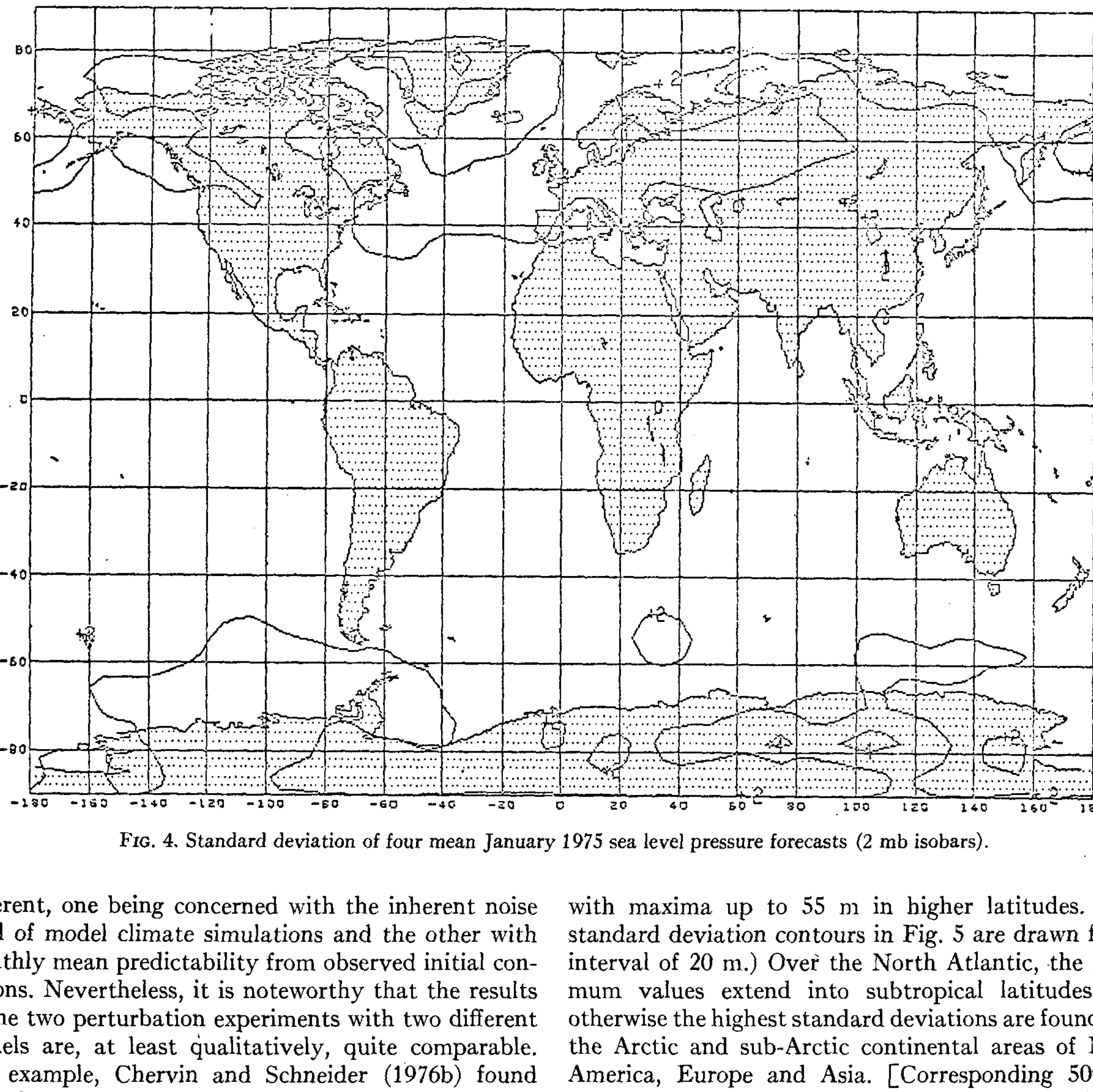

FIG. 4. Standard deviation of four mean January 1975 sea level pressure forecasts (2 mb isobars).

different, one being concerned with the inherent noise level of model climate simulations and the other with monthly mean predictability from observed initial conditions. Nevertheless, it is noteworthy that the results of the two perturbation experiments with two different models are, at least qualitatively, quite comparable. For example, Chervin and Schneider (1976b) found that the NCAR model was relatively insensitive in the tropics to random perturbations, and most sensitive in high latitudes, especially over northern Europe and in the Gulf of Alaska. The GISS model exhibits a similar pattern of response to initial state errors, as shown in Figs. 4–6.

In Fig. 4, maximum values of sea level pressure standard deviations >5 mb, are found in the Bering Sea, with secondary maxima >4 mb over northern Europe. Throughout the tropics, on the other hand, the sea level pressure standard deviations are close to zero, and the values are generally less than 2 mb between latitudes 40°N and 40°S, in good agreement with the results of Chervin and Schneider (1976b). Isobars of sea level pressure standard deviation are drawn for an interval of 2 mb in Fig. 4.

The standard deviations of 500 mb geopotential height, shown in Fig. 5, exhibit a similar pattern of minimum values, generally less than 10 m, in the tropics, with maxima up to 55 m in higher latitudes. (The standard deviation contours in Fig. 5 are drawn for an interval of 20 m.) Over the North Atlantic, the maximum values extend into subtropical latitudes, but otherwise the highest standard deviations are found over the Arctic and sub-Arctic continental areas of North America, Europe and Asia. [Corresponding 500 mb maps are not available for comparison in Chervin and Schneider (1976b).]

The standard deviations of 850 mb temperatures shown in Fig. 6 are comparable with the 1.5 km temperature standard deviations computed with the NCAR nmdel by Chervin and Schneider (1976b). Here the standard deviation isotherms are drawn for an interval of 0.5 K. In the tropics values are generally less than 1K, while in high latitudes of the Northern Hemisphere standard deviations >2 K are found mainly over the continents, with maxima >3 K over North America. However, Fig. 6 also shows values as high as 2 K over the subtropical regions of the North Atlantic and North Pacific Oceans, features not found in the simulations of Chervin and Schneider.

As a further illustration of the relatively modest influence of random initial errors on forecast quality, the standard deviations in Figs. 4–6 may be compared with the absolute errors of the monthly mean control fore-

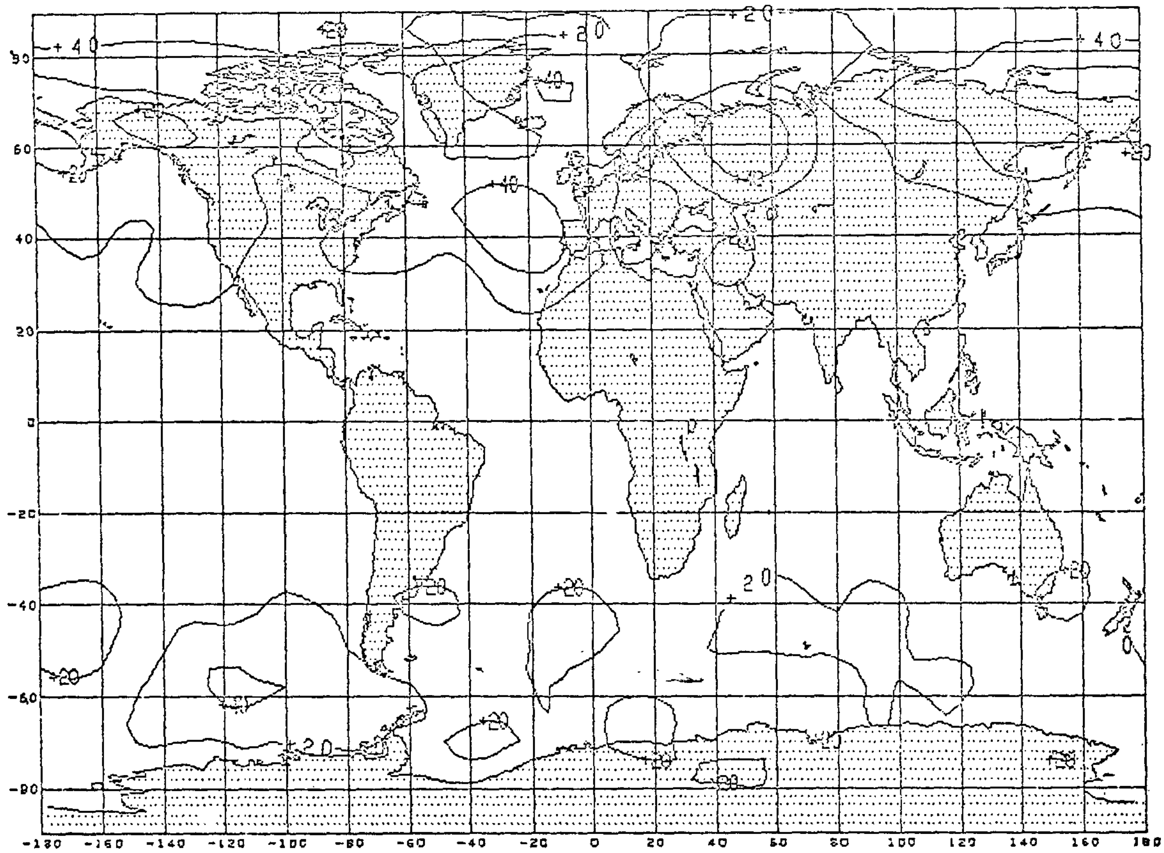

FIG. 5. Standard deviation of four mean January 1975 500 mb height forecasts (20 m contours).

cast. For reasons of brevity, the error maps corresponding to Figs. 4–6 have not been reproduced. However, the major error centers over the Northern Hemisphere are easily identified.

In the sea level pressure forecast, there are nine major error centers over the Northern Hemisphere, ranging from 9 mb over Africa and India, to 11–13 mb over the Pacific, Europe and North America, and to maxima of 14–16 mb over southeast Asia, the Mediterranean, England and northern Russia. In these areas the standard deviations vary from 1–5 mb.

The 500 mb error field for the Northern Hemisphere shows seven major error centers, from 92–97 m over the midwestern Pacific and North America, to 145–160 m over the Atlantic and Mediterranean regions, and with maxima of 202 m over the eastern Pacific and northeast Asia. The standard deviations in these areas range from 10–40 m.

Finally, the 850 mb temperature field for the Northern Hemisphere exhibits eight major errors centers: 7–9 K over Africa, the mid-Pacific, southern Europe, North America, the northwest Pacific and India; and 11–12 K over southeast Asia and the North Atlantic. In these regions the standard deviations vary from 0–3 K.

It is apparent that, while there are significant differences among the four forecasts resulting from the random differences in initial conditions, these are relatively small compared with the errors of the forecasts.

4. Conclusions

Several synoptically significant differences are found among the four monthly mean sea level pressure and 500 mb height fields computed from four slightly different sets of initial conditions for January 1975. However, random errors in the initial state do not appear to represent the major source of large-scale forecast error. All four forecasts exhibit similar types and patterns of error, and similar rms errors and S_1 skill scores. Differences among the forecasts, as measured by these statistics and illustrated by maps of standard deviations of forecast quantities are relatively small compared with the forecast errors. Apparently, the major sources of large-scale forecast error in the monthly mean prognostic maps are not the inevitable random small-scale errors in initial conditions but either unknown systematic large-scale errors in the initial analyses or defects in the model itself.

Acknowledgments. This study was conducted at the Goddard Institute for Space Studies (GISS) in New York City, Robert Jastrow, Director. We are grateful

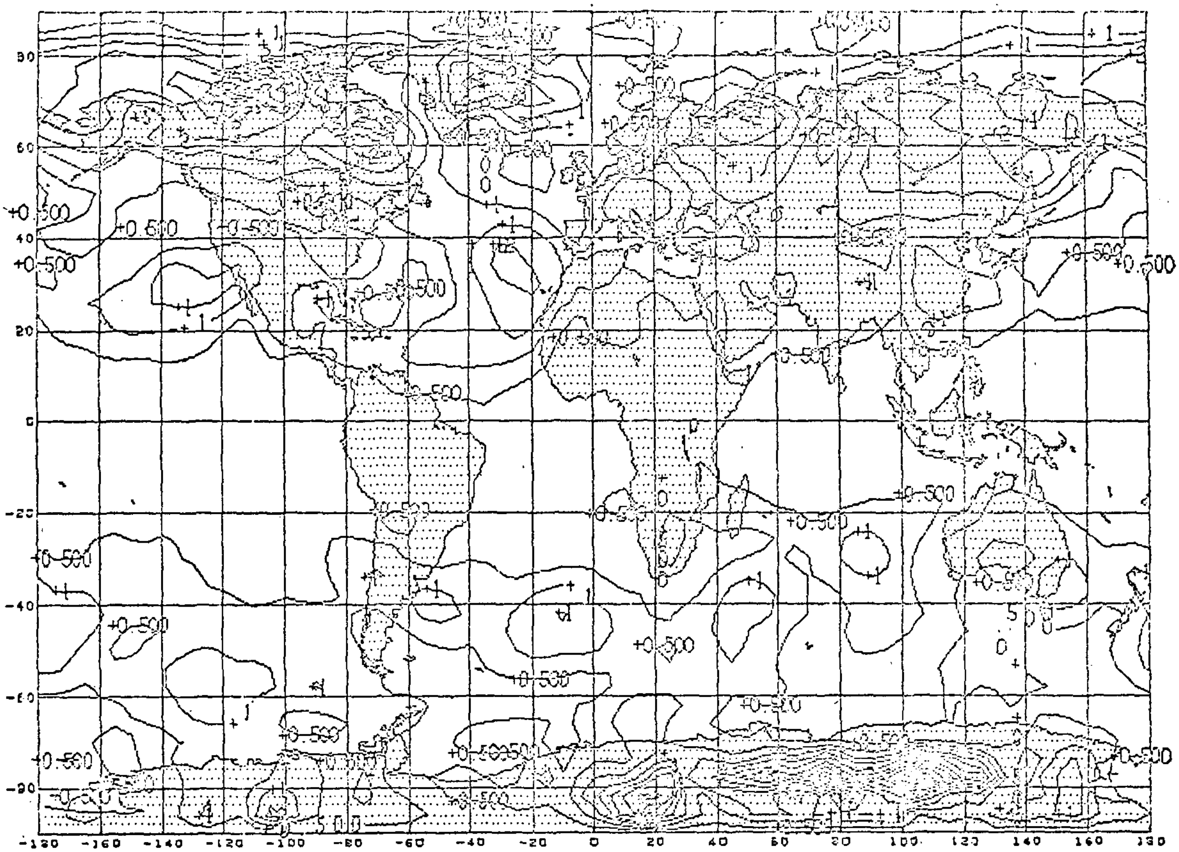

FIG. 6. Standard deviation of four-mean January 1975 850 mb temperature forecasts (0.5 K isotherms).

to the staff of GISS, and especially to Milton Halem, director of the GARP meteorology program there, for providing space, computing facilities, technical assistance and ancillary services. The National Weather Service furnished the global data set. Winthrop T. Johnson and Robert Lutz of the City College also contributed to the study.

REFERENCES

Chervin, R. M., and S. H. Schneider, 1976a: A study of the response of NCAR GGM climatological statistics to random perturbations: estimating noise levels. *J. Atmos. Sci.*, **33**, 391–404.

——, and ——, 1976b: on determining the statistical significance of climate experiments with general circulation models. *J. Atmos. Sci.*, **33**, 405–412.

Druyan, L. M., R. C. J. Somerville and W. J. Quirk, 1975: Extended-range forecasts with the GISS model of the global atmosphere. *Mon. Wea. Rev.*, **103**, 779–795.

Miyakoda, K., G. D. Hembree, R. F. Strickler and I. Shulman, 1972: Cumulative results of extended forecast experiments. I. Model performance for winter cases. *Mon. Wea. Rev.*, **100**, 836–855.

National Academy of Sciences, 1966: The feasibility of a global observation and analysis experiment. Publ. 1290, NAS-NRC, Washington, D. C., 172 pp.

Smagorinsky, J., 1969: Problems and promises of deterministic extended range forecasting. *Bull. Amer. Meteor. Soc.*, **50**, 286–311.

Somerville, R. C. J., P. H. Stone, M. Halem, J. E. Hansen, J. S. Hogan, L. M. Druyan, G. Russell, A. A. Lacis, W. J. Quirk and J. Tenenbaum, 1974: The GISS model of the global atmosphere. *J. Atmos. Sci.*, **31**, 84–117.

Spar, J., 1977: Monthly mean forecast experiments with the GISS model: Correction. *Mon. Wea. Rev.*, **105**, 535–539.

——, R. Atlas and E. Kuo, 1976: Monthly mean forecast experiments with the GISS model. *Mon. Wea. Rev.*, **104**, 1215–1241.

Tenenbaum, J., 1976: Spectral and spatial energetics of the GISS model atmosphere. *Mon. Wea. Rev.*, **104**, 15–30.

Teweles, S., and H. B. Wobus, 1954: Verification of prognostic charts. *Bull. Amer. Meteor. Soc.*, **35**, 455–463.

Vanderman, L. W., J. F. Andrews and J. F. O'Connor, 1976: Extended period forecasting with a global three-layer primitive equation model. *Nat. Wea. Digest*, **1**, 13–24.